

基于激光诱导击穿光谱的页岩与砂岩多元素定量分析

时铭鑫¹, 吴坚^{1*}, 吴迪¹, 李京徽¹, 邱岩², 郭歆宇¹, 周颖¹, 宋自远¹, 陈民鑫¹¹西安交通大学电气工程学院, 电气绝缘与电力设备国家重点实验室, 陕西 西安 710049;²西安交通大学电子科学与工程学院, 电子物理与器件教育部重点实验室, 陕西 西安 710049

摘要 面向天然页岩与砂岩矿物元素组分快速定量分析的检测需求, 笔者搭建了具有自动化分析功能的显微分析激光诱导击穿光谱系统, 并采用该系统对油气页岩与砂岩天然岩石样本进行了快速检测。分析了砂岩与页岩光谱的不确定度差异, 形成了流程化的包括光谱数据提取、筛选、校正、归一化的预处理方法, 有效降低了单一样本与样本间光谱不确定度。基于偏最小二乘回归, 形成了模型输入参数优化流程, 避免了模型过拟合与欠拟合, 提高了定量预测准确度。对岩石中的 Si、Ca、Fe、Al、Mg 等元素进行了定量分析, 多数样本的平均预测均方根误差小于 1%。本研究为天然岩石的矿物组分分析与岩性识别提供了技术支持。

关键词 光谱学; 激光诱导击穿光谱; 页岩; 砂岩; 光谱预处理; 元素定量分析

中图分类号 O436 文献标志码 A

DOI: 10.3788/CJL231332

1 引言

当前, 全球油气勘探开发已进入常规油气稳定发展和非常规油气快速发展阶段。页岩油是一种非常重要的非常规油气资源, 已成为一大开发热点。在油气资源勘探过程中, 获取岩心矿物组分对于岩石岩性识别、储层评估以及油气资源开采具有重要意义^[1-2]。目前, 对岩石矿物组分进行分析的方法包括 X 射线荧光 (XRF)^[3]、电感耦合等离子体发射光谱/质谱法 (ICP-OES/MS)^[4] 以及中子活化分析 (INAA) 等, 这些方法依赖于实验室检测且需要对岩石进行制样与预处理, 会破坏岩石结构而且耗时较长, 难以满足孔隙度快速变化的岩心的检测需求。

激光诱导击穿光谱 (LIBS) 是一种新型的材料元素组分检测技术, 其工作原理为: 聚焦脉冲激光至待测样品表面产生等离子体, 对等离子体的自发光进行光谱分析, 获得样品中元素的种类与含量^[5]。LIBS 技术具有无需样本预处理、多元素同步检测、快速检测、准无损检测等优势, 在冶金分析、地质环境检测、生物医学、核电安全等领域得到了广泛应用^[6-14]。因此, LIBS 技术为油气资源勘探与岩石矿物组分检测提供了可行方案。

面向岩石样本的多元素检测, 国内外研究人员已经取得了一定研究进展。目前已经实现了岩石样本的多元素检测, 如可对 Si、Ca、Mg、Al 等矿物元素进行成

功识别和定量分析, 同时, 基于 LIBS 技术实现了岩石表面二维元素分布的表征^[15-16]。2015 年, Washburn^[17] 将 LIBS 技术应用于油气页岩的检测分析, 结合偏最小二乘 (PLS) 方法初步实现了页岩中多元素定量预测。四川大学的 Xu 等^[18-19] 研究了页岩等离子体温度、电子密度与发射光谱的时间演化过程, 并基于等离子体特性优化实验参数实现了元素定量分析与二维表征。2018 年, Jain 等^[20] 基于 LIBS 技术对开采的页岩样本表面进行了 8 mm × 8 mm 二维多元素绘相, 并将其应用于井场油气资源的识别与储层评价分析。上述研究证明了 LIBS 技术在岩石组分检测方面具有较大的应用优势, 但由于岩石样本矿物组分构成多样并且基质效应显著, LIBS 检测光谱存在不确定度较高、检测准确度和精密度较差的问题。针对这些问题, 研究人员通过改进制样方法与引入机器学习方法降低了光谱不确定性, 提高了定量分析的准确度^[21-23]。然而, 较为复杂的制样方法会增加 LIBS 检测的时间成本, 降低 LIBS 技术的优势, 同时多变量方法的应用尚未形成标准流程, 且极易面临模型过拟合的风险。

针对上述问题, 笔者基于显微分析 LIBS 系统, 对油气页岩与砂岩天然样本进行了多元素定量分析, 形成了流程化的光谱数据提取与预处理方法, 简化了样本预处理过程并降低了光谱数据的不确定性, 提出了模型输入参数优化流程, 在防止模型过拟合的同时提高了模型预测准确度, 为岩石样本多元素高精度检测

收稿日期: 2023-10-27; 修回日期: 2023-11-29; 录用日期: 2023-12-11; 网络首发日期: 2023-12-12

基金项目: 国家重点研发计划青年科学家项目 (SQ2021YFB3703200)、中核集团青年英才项目、国家自然科学基金 (62205266)、中国博士后科学基金 (2022M2514)

通信作者: *jxjawj@mail.xjtu.edu.cn

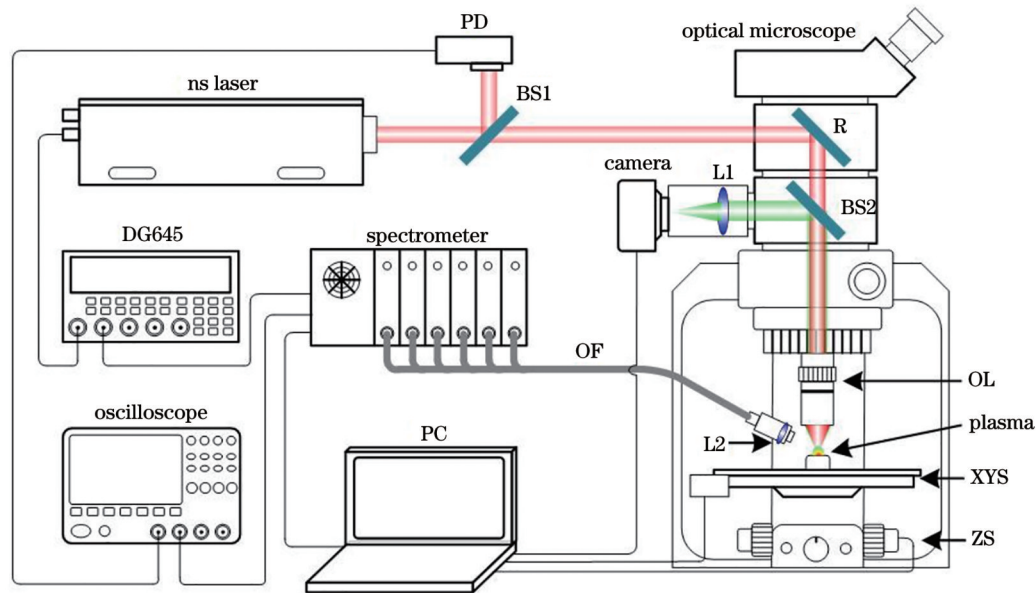
标准的制定提供了参考。

2 实验与方法

2.1 实验设置

实验中使用自主设计搭建的显微分析 LIBS 系统,如图 1 所示,该系统主体由激光器、光谱仪、时序同步器、计算机与光学显微镜组成。使用一台纳秒脉冲激光器(DAWA-100, 镭宝光电)产生激光束(波长为 1064 nm,脉宽为 8 ns,脉冲能量在 0~200 mJ 可调),激光束经过激光分束镜 1(BSF10-C, Thorlabs)分光后由光电二极管(DET10A, Thorlabs)进行采样,以检测激光时序;透射激光束进入一台光学显微镜(BX53M,

Olympus)并通过激光反射镜(NB1-K14, Thorlabs)调节至与显微镜照明光路共轴,最终经过近红外物镜(M Plan APO NIR 20 \times , 焦距 $f=10$ mm, 工作距离 $d_w=20$ mm, Mitutoyo)聚焦于样品表面产生等离子体。等离子体自发光由一短焦透镜收集至光纤并被送入六通道光谱仪(AvaSpec-Mini4096CL-6-Cus, Avantes, 光谱探测范围为 180~885 nm, 平均光谱分辨率为 0.05~0.18 nm)进行光谱采集,积分时间设置为 1 ms。激光器和光谱仪的相对工作时序通过数字延时发生器(DG645, Stanford Research Systems)进行同步,通过计算机软件(DeltaPix Insight)识别、监视相机图像。通过调节 Z 轴电动高度升降台实现自动对焦。



PD: photon detector; DG645: digital delay generator 645; PC: personal computer; BS: beam splitter; R: reflector; L: lens; OL: objective lens; OF: optical fiber; XYS: XY axis electric moving stage; ZS: Z axis electric moving stage

图 1 显微 LIBS 系统设置图

Fig. 1 A diagram for the micro-LIBS system setup

在实验中,调节激光脉冲能量为 30 mJ,光谱仪探测延时设置为 800 ns,通过调节 Z 轴电动高度升降台控制激光聚焦光斑的直径约为 120 μm ,每一个采集点进行 2 发次激光预烧蚀,并累计采集 10 发次光谱。为了降低由样本表面不均匀和等离子体不确定性造成的光谱波动,使用 XY 电动位移平台,在每个样品表面均匀采集 30 幅光谱,并对每三幅光谱进行平均。每个样品最终获取 10 幅光谱。

2.2 样品制备

实验中使用的样品为油气页岩与砂岩天然柱塞岩心与岩屑,共 45 块,其中页岩样本 23 块,砂岩样本 22 块。使用岩石线切割机对所有岩石样本进行切割,以获取适合 LIBS 检测的平整表面。典型的页岩与砂岩照片如图 2 所示。依据国家标准 GB/T 21114—2019《耐火材料 X 射线荧光光谱化学分析 熔铸玻璃片法》,对所有样本中 Si、Ca、Fe、Al、Mg 元素组分参考值

进行测定,所使用的岩石样本中各元素质量分数的平均值、中位数、最小值、最大值、标准偏差(STD)与相对标准偏差(RSD)汇总于表 1。

2.3 分析方法与指标

使用偏最小二乘回归(PLSR)对数据集进行分析处理。偏最小二乘回归是一种经典的多元线性回归模型,它通过在自变量集和因变量集中分别提取主成分来建立回归方程。需要注意的是,输入数据集的大小与最终选择的主成分(PC)的数量会直接影响模型的准确性。因此,在使用模型时通常需要考虑交叉验证的组合,以保证模型预测的有效性。偏最小二乘回归可以通过从高维变量中提取主成分生成一个低维空间来处理高维任务。

为了评估模型的泛化能力,笔者使用留一法交叉验证(LOOCV)对模型进行评估。对于含有 n 个样本的数据集,留一法每次从中选取一个样本作为测试集,

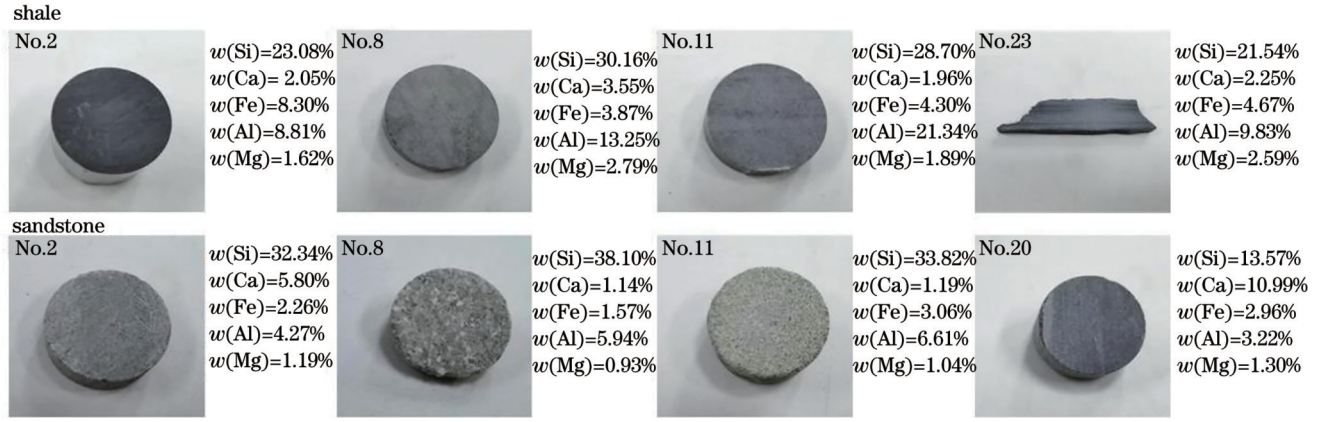


图 2 页岩与砂岩代表性样本照片

Fig. 2 Photos of representative shale and sandstone samples

表 1 页岩与砂岩样本中各元素的质量分数

Table 1 Mass fraction of each element in shale and sandstone samples

Data set	Mass fraction / %					
	Si	Ca	Fe	Al	Mg	
Shale (sample size is 23)	Mean	26.51	2.77	5.61	12.92	2.13
	Median	27.28	2.75	4.67	10.03	2.07
	Min	17.15	0.92	0.79	4.20	0.38
	Max	35.09	5.28	16.67	37.10	5.33
	STD	4.49	1.16	3.42	7.43	1.04
	RSD / %	16.93	41.77	61.13	57.48	49.06
Sandstone (sample size is 22)	Mean	29.06	6.67	2.29	4.61	1.67
	Median	31.43	6.54	2.20	4.49	1.23
	Min	13.57	0.94	1.54	2.69	0.55
	Max	41.02	17.21	3.61	6.61	5.66
	STD	7.39	4.54	0.67	1.18	1.15
	RSD / %	25.42	68.02	29.07	25.69	69.12

将剩余的 $n-1$ 个样本作为训练集,对目标样本进行预测。共循环 n 次完成对所有样本数据的利用。

使用决定系数 (R^2)、训练集均方根误差 (RMSEC, 在公式中记为 ξ_{RMSEC})、预测集均方根误差 (RMSEP, 在公式中记为 ξ_{RMSEP}) 和交叉验证均方根误差 (RMSECV, 在公式中记为 ξ_{RMSECV}) 对模型定量分析结果进行评价。上述评价指标的计算公式为

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i^{\text{Pred}} - y_i^{\text{Ref}})^2}{\sum_{i=1}^n (y_i^{\text{Ref}} - \overline{y_i^{\text{Ref}}})^2}, \quad (1)$$

$$\xi_{\text{RMSEC}} = \sqrt{\frac{\sum_{i=1}^n (y_i^{\text{Pred}} - y_i^{\text{Ref}})^2}{n}}, \quad (2)$$

$$\xi_{\text{RMSEP}} = \sqrt{\frac{\sum_{i=1}^m (y_i^{\text{Pred}} - y_i^{\text{Ref}})^2}{m}}, \quad (3)$$

$$\xi_{\text{RMSECV}} = \sqrt{\frac{\sum_{i=1}^{N_{\text{CV}}} (y_i^{\text{Pred}} - y_i^{\text{Ref}})^2}{N_{\text{CV}}}}, \quad (4)$$

式中: n 为训练集样本数; m 为预测集样本数; N_{CV} 为交叉验证样本数; y_i^{Pred} 与 y_i^{Ref} 为第 i 个样本的预测质量分数与参考质量分数; $\overline{y_i^{\text{Ref}}}$ 为参考质量分数的平均值。

3 分析与讨论

3.1 光谱处理与分析

收集的典型页岩与砂岩的原始光谱及对应的显微照片如图 3 所示。天然岩心样本矿物组成多样、含量变化极大、组分分布不均,导致检测得到的光谱存在较为明显的基质间差异性,特征谱线较多,而且不同样本、同一样本光谱信号波动显著。因此,在定量分析前,需要对所提取的光谱进行谱线预处理、特征谱线识别与分析。

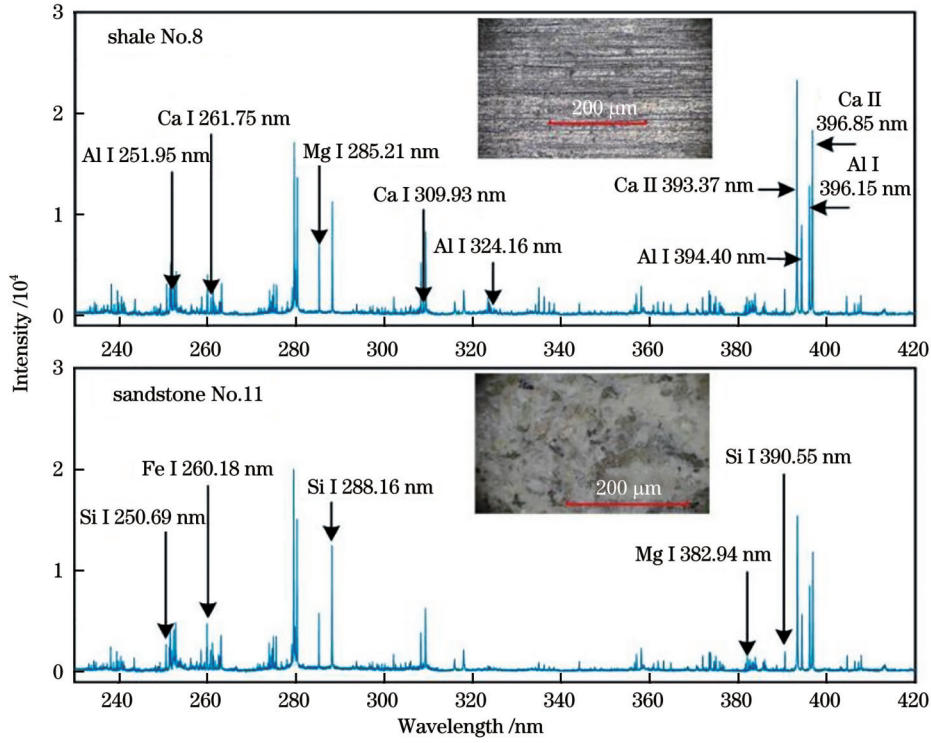


图 3 页岩与砂岩样本的典型光谱

Fig. 3 Typical spectra of shale and sandstone samples

基于 NIST 数据库,对页岩与砂岩中主要元素 Si、Fe、Ca、Al、Mg、Mn、Na、K 的原子谱线、离子谱线进行识别,共识别出 Si 特征原子谱线 14 条、Fe 特征原子谱线 68 条、Fe 特征离子谱线 5 条、Ca 特征原子谱线 32 条、Ca 特征离子谱线 6 条、Al 特征原子谱线 13 条、Al 特征离子谱线 2 条、Mg 特征原子谱线 15 条、Mg 特征离子谱线 3 条。

对所有样本的特征元素光谱进行识别,在谱线中心波长附近±2 个像素内寻找光谱强度最大值并将其作为谱线峰值,然后将峰值附近±5 个像素点储存起来作为输入变量集。天然岩石样本表面元素含量分布不均一性、激光能量发次不确定性、等离子体时空演化不确定性等导致同一样本各采集点光谱间的相对标准偏差较高,其中页岩各元素特征谱线强度的相对标准偏差为 32.50%,砂岩为 56.01%。较高的相对标准偏差会导致定量分析结果的不确定性增大,因此需要对光谱数据集进行预处理。

笔者对光谱作以下预处理:1)基于非二次成本函数对光谱仪每个通道的光谱分别进行基线拟合与校正;2)基于所有特征元素谱线获得输入变量集,然后减去各通道背景值,以提取净强度;3)识别和剔除小于三倍背景标准差(3δ)的谱线信息;4)按光谱仪通道对光谱进行归一化,识别并计算各通道所有符合三倍背景标准差原则的谱线平均峰值,将输入变量集的各谱线峰值除以各通道的平均峰值,得到标准化光谱 I_i^{Norm} ,即

$$I_i^{Norm} = l \times \frac{I_i}{\sum_{j=1}^l I_j}, \quad (5)$$

式中, I_i 为待归一化的谱线峰值, I_j 为该通道中基于三倍背景标准差准则识别到的第 j 条谱线的峰值。

基于预处理前后的光谱数据集分别计算单一样本光谱间的相对标准偏差 R_{single} 以及数据集中各样本光谱间的相对标准偏差 R_{s-s} , 计算公式为

$$R_{single} = \sum_{j=1}^h \frac{\sum_{i=1}^l (I_{ij}^k - \bar{I}_j^k)^2}{h \times l \times \bar{I}_j^k}, \quad (6)$$

$$R_{s-s} = \frac{\sum_{j=1}^h \sum_{k=1}^N (\bar{I}_j^k - \bar{I}_j)^2}{h \times N \times \bar{I}_j}, \quad (7)$$

式中: I_{ij}^k 为第 k 个样本第 i 幅光谱中第 j 个特征谱线的峰值强度; \bar{I}_j^k 为该样本所有光谱第 j 个特征谱线的平均峰值强度; \bar{I}_j 为 k 个样本的第 j 条谱线的平均峰值; l 为同一样本的光谱数量; h 为对应元素的特征谱线数量; N 为样本数量。根据式(6)与式(7)计算 R_{single} 与 R_{s-s} 并将计算结果汇总于表 2。

由表 2 可以看出:经过光谱预处理后,页岩样本单一样本光谱间相对标准偏差(R_{single})的平均值从预处理之前的 32.50% 降低至 14.53%,砂岩样本的平均 R_{single} 从 56.01% 降低至 33.92%;页岩和砂岩样本间光谱相对标准偏差 R_{s-s} 的平均值经光谱预处理后也得到了降低。这证明光谱预处理流程能显著降低光谱不确定性,有助于提高定量分析预测结果的精度。另外,页岩样本各元素的 R_{single} 与 R_{s-s} 显著低于砂岩,这是由岩石

表 2 预处理前后光谱的相对标准偏差
Table 2 Relative standard deviation of spectra before and after pretreatment

unit: %

Data set		Si	Ca	Fe	Al	Mg
Shale	R_{single}^{Origin}	21.00±6.39	57.27±27.04	29.90±16.49	26.15±11.78	28.16±18.22
	R_{S-S}^{Origin}	27.51	63.54	51.44	27.17	40.13
	$R_{single}^{Pretreated}$	10.70±5.15	10.69±17.66	20.99±10.37	14.52±6.71	15.77±7.63
	$R_{S-S}^{Pretreated}$	24.83	61.56	52.58	26.19	37.89
Sandstone	R_{single}^{Origin}	20.52±8.78	88.06±63.69	72.98±29.58	44.30±15.51	54.18±22.18
	R_{S-S}^{Origin}	26.22	81.11	58.69	29.86	56.10
	$R_{single}^{Pretreated}$	7.83±3.22	44.31±28.02	42.49±24.95	34.30±21.73	40.67±16.69
	$R_{S-S}^{Pretreated}$	15.22	73.79	58.69	31.35	53.21

矿物组分差异与颗粒度差异引起的。页岩岩石颗粒粒径与孔隙度均较小,质地相对均匀,光谱不确定度较低;而砂岩组成矿物差异度较大,且粒径与孔隙度较大,带来了较高的光谱不确定性,更依赖于光谱数据预处理。

3.2 多元素组分定量分析

光谱经过预处理后得到 3281 维光谱数据集(作为输入变量),然而,较高的数据集维数和较高的主成分会导致模型出现过拟合,从而导致模型对训练样本预测结果较好而对新预测样本的预测结果较差,因此需要对模型输入数据维度与主成分进行优化。笔者建立了流程化的模型输入参数的优化流程,具体包括:1)识别特征元素谱线峰值及其附近±2个像素的光谱数据;2)选取具有最高相关系数的*i*×5个光谱像素点作为输入变量;3)选定主成分数为*j*;4)计算得到每次交叉验证对应的 ξ_{RMSECV} 以及 ξ_{RMSEC} 的平均值;5)改变*i*和*j*,遍历所有情况,寻找 $\sqrt{\xi_{RMSECV}^2 + \xi_{RMSEC}^2}$ 最小的输入参数并将其作为优化结果。模型的优化流程如图 4 所示。

本文取*i*最小值为 1,最大值为 40,即输入光谱维度最大为 200,输入数据主成分(PC)的值最大为 12,优化过程如图 5 所示。

以 Si 元素为例,图 5 展示了优化参数随输出光像素数、主成分数量的变化。当输入光谱维度固定时,优化参数随主成分数增加呈现先降低后增加的趋势。这是由于较少的主成分数会引起模型欠拟合,训练集样本解释不足, RMSEC 偏高,而过大的主成分数则会引起模型过拟合,对新预测集的解释较差, RMSECV 值偏高。只有选择合适的主成分数,才能使 PLSR 模型得到最佳性能。优化参数完成后,得到各元素最佳适配主成分数与输入光谱。图 6 展示了 Si 的最优输入光谱,其主要包含 Si 特征原子谱线及其附近的光谱信息。

图 7 展示了在最佳输入光谱与主成分下各元素的

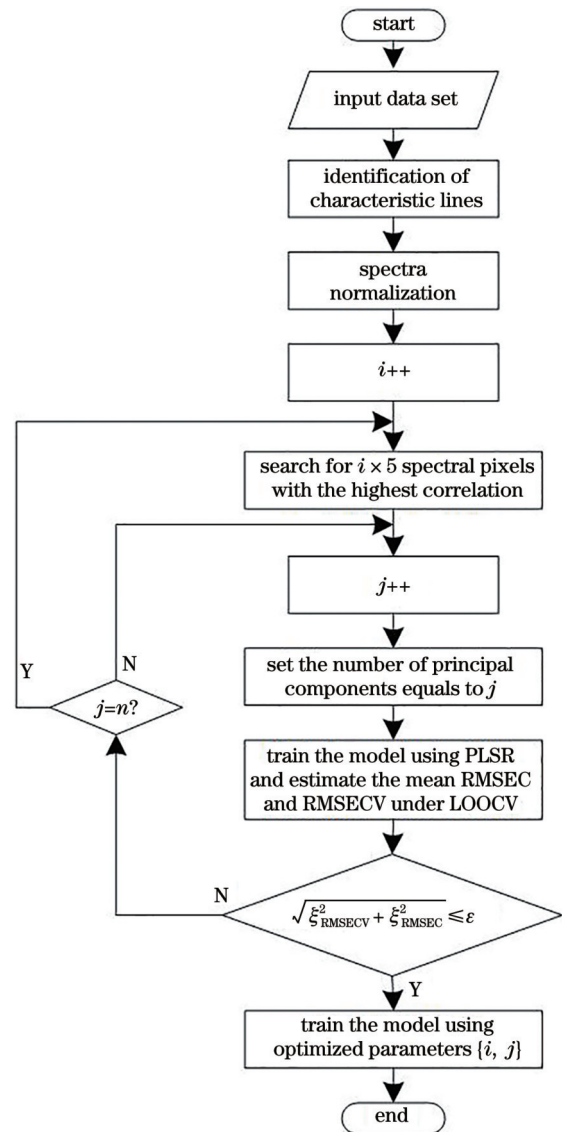


图 4 PLSR 参数优化流程图

Fig. 4 PLSR parameter optimization flow chart

预测结果(从页岩与砂岩样本中各选取 5 个作为测试集,剩余的 35 个作为训练集,对两种岩石分别进行训练),PLSR 对各元素训练结果的评价指标汇总于表 3。

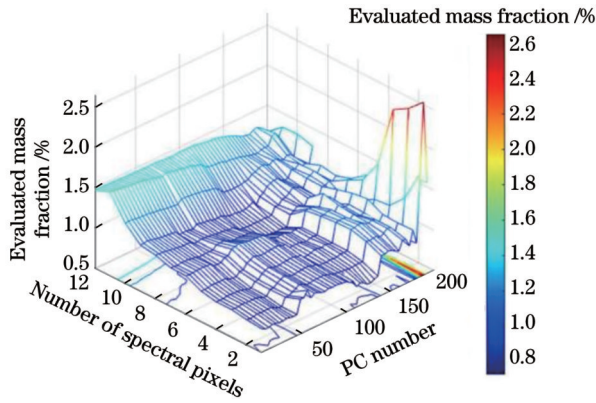


图 5 Si 元素的参数优化过程

Fig. 5 Parameter optimization process of Si element

各元素质量分数预测值与参考值具有较高的线性拟合度,大部分元素的 $R^2 > 0.95$, RMSEC 小于 1%, 说明 PLSR 模型可以较好地实现页岩与砂岩的多元素定

量分析,同时大部分元素的 RMSECV 小于 2%, 说明训练集模型交叉验证结果良好,极大地避免了过拟合现象。对于模型测试集,大部分元素的 RMSEP 小于 1%, 说明 PLSR 模型对页岩与砂岩中主要元素的预测准确度较高,其中砂岩中的 Si 元素与页岩中的 Al 元素存在相对较高的 RMSEP, 分别为 2.6230% 与 2.8818%。这是由元素较大的含量区间跨度以及较高含量下谱线的自吸收现象共同造成的。分析图 7 可以得出砂岩样本的预测 RSD(即图中的误差棒)明显高于页岩样本,这与表 3 所示结果一致,证明了砂岩表面矿物元素均存在更高的不均一性。另外,砂岩与页岩存在明显的含量区间差异性,其中砂岩中的 Si、Ca、Mg 元素存在较大的含量差异,而页岩中的 Fe、Al 元素存在较大的含量差异。更重要的是,两种岩石中的元素含量分布存在较少的交叉与重合,这为岩石岩性识别与分类提供了重要的数据支持。

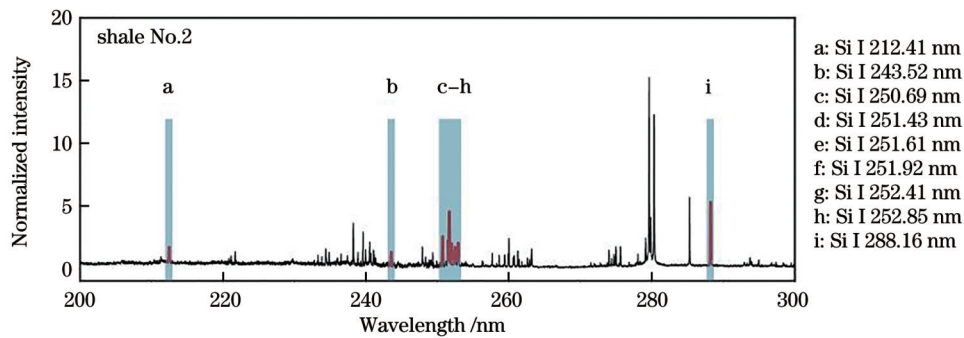


图 6 参数优化后的输入光谱(页岩样本 2)

Fig. 6 Input spectrum after parameter optimization (shale No. 2)

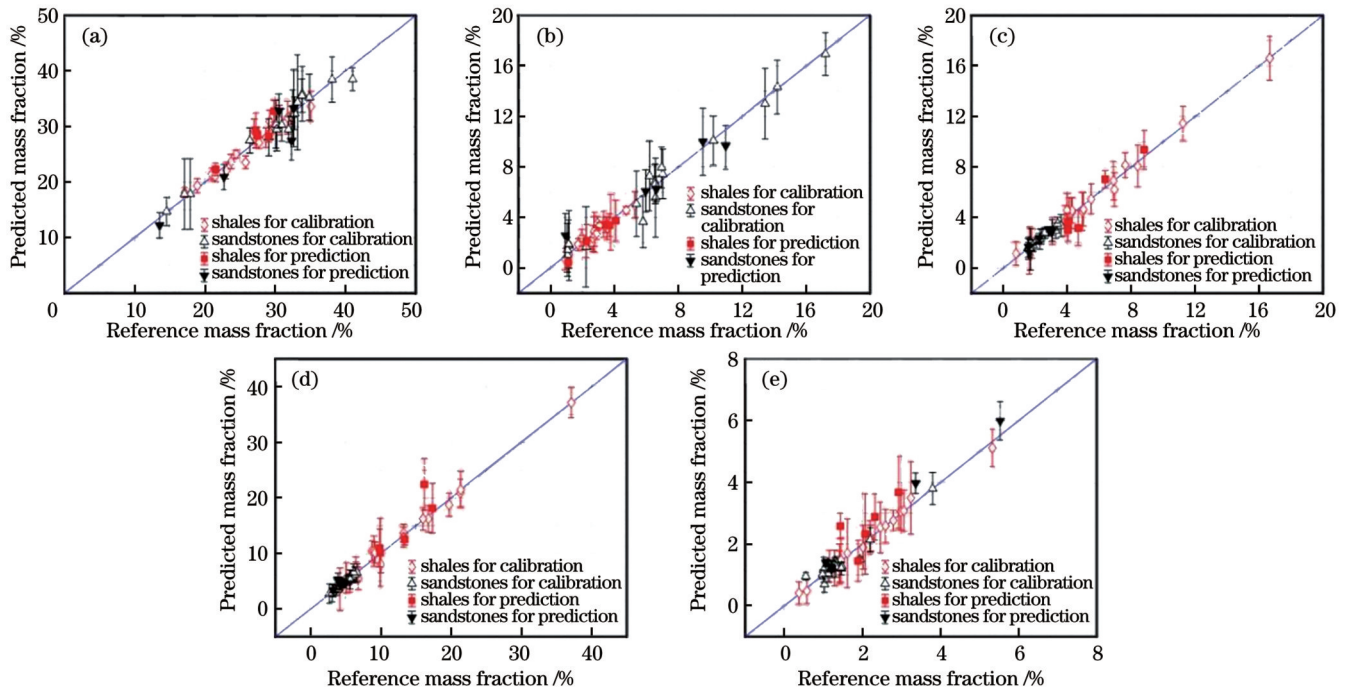


图 7 Si、Ca、Fe、Al、Mg 元素的定标曲线。(a) Si 元素;(b) Ca 元素;(c) Fe 元素;(d) Al 元素;(e) Mg 元素

Fig. 7 Calibration curves of Si, Ca, Fe, Al, and Mg elements. (a) Si element; (b) Ca element; (c) Fe element; (d) Al element; (e) Mg element

表 3 Si、Ca、Fe、Al、Mg 元素定量分析结果的评价指标
Table 3 Evaluation indexes of quantitative analysis of Si, Ca, Fe, Al, Mg elements

Sample and element	R^2	RMSEC / %	RMSEP / %	RMSECV / %	
Shale	Si	0.9524	1.0270	1.7494	1.9027
	Ca	0.9170	0.3286	0.3615	0.7467
	Fe	0.9915	0.3369	0.9181	1.0515
	Al	0.9870	0.9177	2.8818	3.8290
	Mg	0.9909	0.1071	0.7705	0.3389
Sandstone	Si	0.9697	1.2171	2.6230	3.4010
	Ca	0.9737	0.7594	0.9745	1.5270
	Fe	0.9457	0.1529	0.1355	0.3406
	Al	0.9661	0.2448	0.7133	0.7810
	Mg	0.9093	0.2090	0.3744	0.3809

4 结 论

面向天然页岩与砂岩矿物样本元素含量快速定量检测的需求,笔者基于显微分析 LIBS 系统,针对岩石元素定量分析准确度低的难题,研究了光谱数据的预处理与模型优化方法,并对油气页岩与砂岩样本进行了多元素定量检测。

针对岩石样本基质复杂、元素分布不均等问题,笔者建立了光谱数据预处理标准流程,通过谱线筛选、基线校正、通道归一化等方法实现了页岩样本光谱间平均相对标准偏差从 32.50% 降低至 14.53%,砂岩样本从 56.01% 降低至 33.92%,显著降低了光谱不确定性。针对岩石样本定量预测准确度低的难题,笔者提出了一种基于 PLSR 的定量分析流程,通过遍历输入谱线维度与主成分数得到最优 RMSEC 与 RMSECV,避免了模型欠拟合与过拟合,实现了页岩与砂岩中 Si、Ca、Fe、Al、Mg 等元素的定量分析,其中多数样本的 $R^2 > 0.95$, RMSEC 小于 1%, RMSEP 小于 1%。本研究为天然油气页岩与砂岩矿物组分的快速分析提供了技术支持,所提出的数据预处理与定量分析流程有望成为 LIBS 分析的技术标准。

参 考 文 献

- Ross D J K, Bustin R M. Investigating the use of sedimentary geochemical proxies for paleoenvironment interpretation of thermally mature organic-rich strata: examples from the Devonian-Mississippian shales, Western Canadian Sedimentary Basin[J]. *Chemical Geology*, 2009, 260(1/2): 1-19.
- Zhao X G, Kang J L, Lan B. Focus on the development of shale gas in China: based on SWOT analysis[J]. *Renewable and Sustainable Energy Reviews*, 2013, 21: 603-613.
- Nikonow W, Rammlair D, Meima J A, et al. Advanced mineral characterization and petrographic analysis by μ -EDXRF, LIBS, HSI and hyperspectral data merging[J]. *Mineralogy and Petrology*, 2019, 113(3): 417-431.
- Momen A A, Zachariadis G A, Anthemidis A N, et al. Use of fractional factorial design for optimization of digestion procedures followed by multi-element determination of essential and non-essential elements in nuts using ICP-OES technique[J]. *Talanta*, 2007, 71(1): 443-451.
- 李祥友, 刘可, 周冉, 等. 激光诱导击穿光谱技术及应用综述[J]. *中国激光*, 2022, 49(12): 1202003.
Li X Y, Liu K, Zhou R, et al. Laser-induced breakdown spectroscopy and its application[J]. *Chinese Journal of Lasers*, 2022, 49(12): 1202003.
- 王金梅, 卢健树, 郑培超, 等. 基于 LIBS 技术结合内标法的烧结矿碱度测量研究[J]. *中国激光*, 2023, 50(22): 175-182.
Wang J M, Lu J S, Zheng P C, et al. Study on alkalinity measurement of sinter based on libs technology combined with internal standard method[J]. *Chinese Journal of Lasers*, 2023, 50(22): 175-182.
- 邱苏玲, 李安, 王宪双, 等. 基于激光诱导击穿光谱的矿石中铁含量的高准确度定量分析[J]. *中国激光*, 2021, 48(16): 1611002.
Qiu S L, Li A, Wang X S, et al. High-accuracy quantitatively analysis of iron content in mineral based on laser-induced breakdown spectroscopy[J]. *Chinese Journal of Lasers*, 2021, 48(16): 1611002.
- 时铭鑫, 王傲松, 黄大鹏, 等. 基于支持向量机和随机森林算法结合光纤式激光诱导击穿光谱定量检测核电用钢中铬[J]. *冶金分析*, 2021, 41(1): 30-40.
Shi M X, Wang A S, Huang D P, et al. Quantitative analysis of chromium by fiber-optic laser-induced breakdown spectroscopy of steels for nuclear power plant based on support vector regression and random forest regression[J]. *Metallurgical Analysis*, 2021, 41(1): 30-40.
- Dyar M D, Giguere S, Carey C, et al. Comparison of baseline removal methods for laser-induced breakdown spectroscopy of geological samples[J]. *Spectrochimica Acta Part B: Atomic Spectroscopy*, 2016, 126: 53-64.
- 白伟洋, 陈蔚芳, 杨成杰, 等. 基于 LIBS 技术的不锈钢精细分类方法[J]. *激光与光电子学进展*, 2022, 59(23): 2330001.
Bai W Y, Chen W F, Yang C J, et al. Fine classification method of stainless steel based on LIBS technology[J]. *Laser & Optoelectronics Progress*, 2022, 59(23): 2330001.
- 沙文, 李江涛, 鲁翠萍. 基于激光诱导击穿光谱技术寻优定量分析土壤中 Mn 元素[J]. *中国激光*, 2020, 47(5): 0511001.
Sha W, Li J T, Lu C P. Quantitative analysis of Mn in soil based on laser-induced breakdown spectroscopy optimization[J]. *Chinese Journal of Lasers*, 2020, 47(5): 0511001.
- 潘超超, 赵南京, 马明俊, 等. 基于激光诱导击穿光谱技术的土壤镉元素高灵敏检测[J]. *激光与光电子学进展*, 2023, 60(17): 1730006.
Pan C C, Zhao N J, Ma M J, et al. Highly sensitive detection of Cd in soil using laser-induced breakdown spectroscopy[J]. *Laser & Optoelectronics Progress*, 2023, 60(17): 1730006.

- [13] 李晨红, 闫薪如, 辛英健, 等. 基于 LIBS 技术结合 AFSA-SVM 算法对岩石类别的识别[J]. 激光与光电子学进展, 2023, 60(9): 0930002.
Li C H, Yan X R, Xin Y J, et al. Rock identification using LIBS technique combined with AFSA-SVM algorithm[J]. Laser & Optoelectronics Progress, 2023, 60(9): 0930002.
- [14] 陈世航, 刘莉, 邓志伟, 等. 基于单激光构造的 DP-LIBS 对钽铌矿的检测研究[J]. 激光与光电子学进展, 2023, 60(13): 1330001.
Chen S H, Liu L, Deng Z W, et al. Detection of tantalum-niobium ores using double-pulse laser-induced breakdown spectroscopy based on a single laser[J]. Laser & Optoelectronics Progress, 2023, 60(13): 1330001.
- [15] Sanghavi H K, Jain J, Bol'shakov A, et al. Determination of elemental composition of shale rocks by laser induced breakdown spectroscopy[J]. Spectrochimica Acta Part B: Atomic Spectroscopy, 2016, 122: 9-14.
- [16] Quarles C D, Jr, Miao T, Poirier L, et al. Elemental mapping and characterization of petroleum-rich rock samples by laser-induced breakdown spectroscopy (LIBS)[J]. Fuels, 2022, 3(2): 353-364.
- [17] Washburn K E. Rapid geochemical and mineralogical characterization of shale by laser-induced breakdown spectroscopy [J]. Organic Geochemistry, 2015, 83/84: 114-117.
- [18] Xu T, Zhang Y, Zhang M, et al. Temporal-resolved characterization of laser-induced plasma for spectrochemical analysis of gas shales[J]. Spectrochimica Acta Part B: Atomic Spectroscopy, 2016, 121: 28-37.
- [19] Xu T, Liu J, Shi Q, et al. Multi-elemental surface mapping and analysis of carbonaceous shale by laser-induced breakdown spectroscopy[J]. Spectrochimica Acta Part B: Atomic Spectroscopy, 2016, 115: 31-39.
- [20] Jain J, Quarles C D, Jr, Moore J, et al. Elemental mapping and geochemical characterization of gas producing shales by laser induced breakdown spectroscopy[J]. Spectrochimica Acta Part B: Atomic Spectroscopy, 2018, 150: 1-8.
- [21] Yao S C, Zhao J B, Xu J L, et al. Optimizing the binder percentage to reduce matrix effects for the LIBS analysis of carbon in coal[J]. Journal of Analytical Atomic Spectrometry, 2017, 32 (4): 766-772.
- [22] Shi L L, Lin Q Y, Duan Y X. A novel specimen-preparing method using epoxy resin as binding material for LIBS analysis of powder samples[J]. Talanta, 2015, 144: 1370-1376.
- [23] 赵文雅, 闵红, 刘曙, 等. 神经网络在激光诱导击穿光谱数据分析中的应用进展[J]. 光谱学与光谱分析, 2021, 41(7): 1998-2004.
Zhao W Y, Min H, Liu S, et al. Application progress of artificial neural network in laser-induced breakdown spectral data analysis [J]. Spectroscopy and Spectral Analysis, 2021, 41(7): 1998-2004.

Multi-Element Quantitative Analysis of Shale and Sandstone Based on Laser Induced Breakdown Spectroscopy

Shi Mingxin¹, Wu Jian^{1*}, Wu Di¹, Li Jinghui¹, Qiu Yan², Guo Xinyu¹, Zhou Ying¹,
Song Ziyuan¹, Chen Minxin¹

¹State Key Laboratory of Electrical Insulation and Power Equipment, School of Electrical Engineering, Xi'an Jiaotong University, Xi'an 710049, Shaanxi, China;

²Ministry of Education Key Laboratory of Electronic Physics and Devices, School of Electronic Science and Engineering, Xi'an Jiaotong University, Xi'an 710049, Shaanxi, China

Abstract

Objective Shale oil is an important unconventional resource that has become a major development focus. Owing to the possible reserves of petroleum and natural gas, information from core samples, which mainly contain geochemical and mineralogical information, has attracted increasing attention from geologists and energy enterprises. Conventional analytical methods for mineralogical analysis, including inductively coupled plasma optical emission or mass spectrometry (ICP-OES/MS), instrumental neutron activation analysis (INNA), and X-ray fluorescence (XRF), are unsuitable for the *in-situ* analysis of shale because of their slow analysis speed and limited detection range. Laser-induced breakdown spectroscopy (LIBS) is a novel analytical method for various materials that uses laser pulses focused on a sample to generate laser-produced plasma and spectral emissions. With the advantages of minimal sample preparation, remote analysis, microsample ablation, and high analytical speed, LIBS is an optional access method for analysis of cores. In this study, multi-element quantitative analysis of natural shale and sandstone samples is conducted using a micro-LIBS system. A standardized spectral preprocessing method is developed to reduce spectral uncertainty, and a model input parameter optimization process is established to prevent model overfitting while improving model prediction accuracy. A quantitative analysis of Si, Ca, Fe, Al, Mg, and other elements in the shales and sandstones is performed, and the average predicted root-mean-square error of most samples is less than 1%, providing technical support for regular mineral component analysis and lithology identification of shales and sandstones.

Methods A micro-LIBS system was established in this study, consisting of a laser, spectrometer, timing synchronizer, computer, and optical microscope. The system used a nanosecond-pulse laser to generate a laser beam that passes through a laser beam splitter and was sampled by a photodiode to detect the laser timing. The transmitted laser beam was introduced into an optical microscope,

adjusted to be coaxial with the microscope illumination light path through a laser mirror, and finally focused on the sample surface through a near-infrared objective lens to generate plasma. Plasma emission was collected by a short-focus lens to an optical fiber and connected to a six-channel spectrometer for spectrum collection; the integration time was set to 1 ms. The laser pulse energy was adjusted to 30 mJ, and the detection delay of the spectrometer was set to 800 ns. Before the spectrum acquisition, two laser shots were applied for pre-ablation to stabilize the pulse energy, and 10 accumulated spectra were collected and analyzed. Partial least squares regression was used for the data analysis.

Results and Discussions Because of the variety, inhomogeneity, and complexity of the samples, it is necessary to perform spectral line preprocessing and characteristic spectral line identification of the extracted spectra before quantitative analysis. Therefore, this study establishes a set of spectral data pre-processing procedures. After spectral preprocessing, the average relative standard deviation of a single shale sample (R_{single}) decreases from 32.50% to 14.53%, whereas the average R_{single} of the sandstone samples decreases from 56.01% to 33.92%. Simultaneously, the spectral relative standard deviation $R_{\text{s-s}}$ between shale and sandstone samples is also reduced after spectral processing, proving that the spectral preprocessing process can significantly reduce spectral uncertainty and improve the precision of the quantitative analysis prediction results. Based on the leave-one-out method, this study constructs a set of PLSR model input dataset parameter optimization processes. When the number of input spectra is fixed, the optimization parameters initially decrease, and then increase as the number of principal components increases. This is because lower principal components cause model underfitting and insufficient interpretation of the training set samples, and the RMSEC is too high. The number of principal components causes overfitting of the model, poor interpretation of the new prediction set, and a high RMSECV value. The PLSR model achieves the best performance only when an appropriate number is selected. Under the optimized parameters, $R^2 > 0.95$ and $\text{RMSEC} < 1\%$ for most samples, indicating that the PLSR model can better achieve multi-element quantitative analysis of shale and sandstone.

Conclusions To meet the demands of rapid quantitative determination of mineral element content in natural shale and sandstone samples, a microanalysis laser-induced breakdown spectroscopy system with automated analysis capabilities is developed to rapidly detect and analyze oil and gas shale and sandstone natural samples. A standard process for spectral data preprocessing is established, and the average relative standard deviation of shales is reduced from 32.50% to 14.53%, and from 56.01% to 33.92% for sandstones. This study proposes a quantitative analysis process based on partial least squares regression (PLSR) with Si, Ca, Fe, Al, and Mg quantitatively predicted. The R^2 value for most samples is 0.95, and the RMSEC and RMSEP values are lower than 1%. This study provides technical support for the rapid analysis of the mineral components in shale and sandstone samples. The proposed data preprocessing and quantitative analysis is expected to become a technical standard for LIBS analyses.

Key words spectroscopy; laser induced breakdown spectroscopy; shale; sandstone; spectral pretreatment; elemental quantification